

Two Forms of Inconsistency in Quantum Foundations

Jeremy Steeger
jsteeger@nd.edu

Nicholas Teh
nteh@nd.edu

September 10, 2018

Abstract

Recently, there has been some discussion of how Dutch Book arguments might be used to demonstrate the rational incoherence of certain hidden variable models of quantum theory (Feintzeig 2015; Feintzeig and Fletcher 2017; Wroński and Godziszewski 2017). In this paper, we argue that the ‘form of inconsistency’ underlying this alleged irrationality is deeply and comprehensively related to the more familiar ‘inconsistency’ phenomenon of contextuality. Our main result is that the hierarchy of contextuality due to Abramsky and Brandenburger (2011) corresponds to a hierarchy of additivity/convexity-violations which yields formal Dutch Books of different strengths. We then use this result to provide a partial assessment of whether these formal Dutch Books can be given a convincing normative interpretation.

1 Introduction

Within the philosophical foundations of quantum theory, there is a long tradition of attempting to argue that the predictions of quantum theory cannot be reproduced by hidden variable theories that assume a classical structure of events, probabilities of events, states of affairs and functional relationships relating these items (call these **CHVTs**). Very roughly speaking, such ‘no-go’ arguments take the following form:

(Standard) There exists a quantum model whose empirical probabilities are inconsistent with the probabilities that can be reproduced by the relevant CHVT.¹

For example, the GHZ, Kochen-Specker, Hardy, and Bell quantum models all provide instances of (Standard) arguments.

The development of this tradition has since resulted in a veritable cornucopia of models witnessing such inconsistencies with classical intuitions. Thus, one important avenue of contemporary research (Abramsky and Brandenburger 2011; Acín et al. 2012; Liang et al. 2011) seeks to impose order on this landscape by uncovering what we shall refer to as ‘forms of inconsistency’, i.e. high-level structures that provide a unifying explanation—and perhaps even a classification—of how these models give rise to such inconsistencies. The goal of this paper is to argue for a thoroughgoing and novel connection between two forms of inconsistency in quantum foundations, and to apply this result to several philosophical puzzles.

The first form, which has a long and distinguished pedigree, is the *Gluing Inconsistency* that is suggested by ‘contextuality arguments’ against CHVTs, as exemplified by the Kochen-Specker theorem (Kochen and Specker 1975). Its structural essence has recently been rigorously articulated by Abramsky and Brandenburger (henceforth **AB**), who not only generalize it to apply to a large class of physical theories, but also provide a three-tiered ‘hierarchy of contextuality’ which classifies the ways in which physical models can deviate from the classical probability assumed by a CHVT.

The second form, which we call *Dutch Bookability*, is somewhat more recent and thus less explored; indeed, much of the work of this paper will be to show that it really is a ‘form of inconsistency’ in our sense. In its simplest incarnation, it arises in (Feintzeig 2015) and (Feintzeig and Fletcher 2017) and appears to differ from the first form in two ways. First, it is not presented as arising from an argument against

¹We will give more detail in Sections 2 and 3 about the class of CHVTs that we wish to consider. In brief: we will only consider *noncontextual* ‘factorizable CHVTs’ in the sense of (Abramsky and Brandenburger 2011, Section 8); thus, (Standard) arguments do not apply to Bohmian mechanics, which is both contextual and non-local.

a CHVT, but rather against a *Generalized HVT (GHVT)*, by which we mean a hidden variable theory based on a weakening of classical probability theory, in a sense that we will make precise below. Second, the ‘inconsistency’ in question stems from the fact that the GHVT is ‘Dutch Bookable’ in the sense that it violates a certain formal constraint. Feintzeig and Fletcher (2017) then make a move that has become standard in philosophical discussions concerning *classical* probability: they interpret this constraint normatively, i.e. as a rational constraint on the expectations of an epistemic agent, and thus take its violation to provide a no-go argument against a subjective interpretation of the probabilities of GHVTs.² For the sake of clarity, we will distinguish between the formal constraint and its normative interpretation by using ‘*Dutch Bookability*’ to refer to the violation of the formal constraint, and ‘normative Dutch Bookability’ to refer to the normative interpretation of this violation.

Despite these apparent differences, we will argue that Gluing Inconsistency and Dutch Bookability are really two ways of understanding the same set of facts which fall under (Standard), because each form allows us to represent the empirical probabilities of a physical model and to classify the ways in which certain HVTs fail to reproduce these probabilities. Our argument culminates in the result (Theorem 1) that, for a large class of physical models (including quantum models), AB’s three-tiered hierarchy of Gluing Inconsistency corresponds to a three-tiered hierarchy of Dutch Bookability, which provides the aforementioned classification.

This result yields powerful applications to two puzzles raised by the recent exploration of the relationship between Dutch Books and quantum theory. First, the only connection between GHVT models of quantum theory and Dutch Bookability that has appeared in the literature (Feintzeig 2015; Feintzeig and Fletcher 2017) stems from a ‘finite null cover’—an incredibly strong (indeed maximal!) violation of ‘subadditivity’ which leads to Dutch Books. What explanation can be given of the strength of this violation? Our main result allows us to answer this question by explaining the strength of the violation in terms of AB’s three-tiered classification of the contextuality exhibited by quantum (and non-quantum) models: the ‘maximal Dutch Bookability’ of the above violation derives from its being the strongest form of contextuality (exemplified in quantum theory by Kochen-Specker models), but there are also quantum models which exhibit weaker forms of Dutch Bookability corresponding to the second and third tier respectively, viz. the Hardy model (Hardy 1993) and the Bell model (Bell 1964).

Second, what room is there to interpret Dutch Bookability normatively, i.e. as a no-go argument against a subjective approach to the probabilities of noncontextual GHVTs? This question is difficult to fully assess because the relevant hidden variable theories have not been clearly defined in the literature. Nonetheless, we will suggest the normative force of Dutch Bookability hinges on an approach to GHVTs that borrows an interpretive strategy from classical probability theory—and this approach is at the very least controversial.

The paper is structured as follows. First, Section 2 provides a non-technical introduction to our topic that is accessible to the general reader. Section 2.1 reviews the notion of ‘contextuality’ that is central to this paper. Section 2.2 introduces the first form of inconsistency, viz. AB’s Gluing Inconsistency, and Section 2.3 introduces the second form of inconsistency, viz. (formal) Dutch Bookability. Section 2.3 then gives a conceptual explanation of the main results of our paper and their motivation. At this point, the general reader will have sufficient background to skip straight ahead to the conclusion (Section 5).

Sections 3 and 4 are aimed more at specialist readers in the philosophy of physics and formal epistemology. However, we stress that, even here, technical prerequisites have been kept to an absolute minimum and our arguments are accessible to a reader with only knowledge of elementary set theory. Thus, a general reader who is willing to work through the details should be able to follow the gist of our arguments.

Section 3.1 provides an elementary review of AB’s highly influential framework for contextuality, which demonstrates that contextuality comes in different strengths or ‘tiers’ of a hierarchy. To the best of our knowledge, this is the first such treatment aimed at philosophers. Section 3.2 then introduces one of our novel contributions, viz. the construction of a translation manual from AB’s framework to the ‘weak probability spaces’ that are used in order to formulate Dutch Bookability. The details of this translation manual are subtle and themselves of conceptual significance, since they describe how to give a quasi-probabilistic interpretation to one of the deepest features of quantum theory, viz. contextuality.

Section 4 draws on the material of Section 3 to prove our main result, viz. the correspondence between AB’s hierarchy for contextuality and a hierarchy of (formal) Dutch Bookability. We begin by situating our

²Feintzeig and Fletcher also argue that violating the formal constraint associated with ‘Dutch Books’ is pathological even when the probabilities in question are given a frequentist interpretation, but we will not address this argument here.

argument with respect to previous work on a hierarchy of Dutch Bookability that was conducted within a purely abstract setting, i.e. that did not relate this hierarchy to physical models (and quantum theory in particular). By contrast, Section 4.1 relates the AB hierarchy—and its connection with *physical models*—to a hierarchy of Dutch Bookability in terms of ‘additivity-violation’, and Section 4.2 does the same for a hierarchy of Dutch Bookability in terms of ‘convexity-violation’. We note that the formal epistemology community will be especially interested in the results of Section 4.2, since the standard modern approach to Dutch Books is by means of convex analysis.³

Finally, Section 5 addresses the question of how formal Dutch Bookability may be interpreted normatively, i.e. as a no-go against a subjective interpretation of the probabilities of GHVTs. We first offer one plausible reconstruction of Feintzeig and Fletcher’s normative Dutch Book argument, and we show that the soundness of this reconstructed argument hinges on a certain choice of response functions, i.e. probabilities of events conditional on certain states of affairs. We then note that a GHVT need not specify response functions in this way.

2 Contextuality and Inconsistency

2.1 Contextuality

The subject of ‘contextuality’ was introduced into quantum foundations in (Bell 1966) and (Kochen and Specker 1975), after which it has become common to characterize quantum theory as ‘contextual’. In order to explain what this means, we now review several preliminary definitions. First, a (measurement) *context* is a set of ‘co-measurable’ or ‘compatible’ measurements, i.e. measurements that can be made jointly (thus, in the particular case of quantum theory, a ‘context’ refers to a set of commuting observables). Second, a *noncontextual* (NC) CHVT is one whose ‘response functions’ (i.e. functions that specify the probabilities of measurement events conditional on the system’s being in some particular state) are independent of any information about contexts.⁴ Third, we define a *possibility assignment* as an assignment of either 1 or 0 to events. Following Spekkens (2005), we will use the term *outcome-deterministic* (OD) to refer to an CHVT whose response functions only make possibility assignments.

The characterization of quantum theory as ‘contextual’ then arises through the following ‘quantum contextuality argument’, which is a special case of (Standard): there exists a quantum model whose empirical possibilities are inconsistent with the response functions of an (NC, OD) CHVT. We say that quantum theory is ‘contextual’ because such quantum models exist.

We now discuss how quantum contextuality suggests a geometric form of inconsistency that we shall call ‘Gluing Inconsistency’. One classic way of demonstrating quantum contextuality is by means of a (discrete) Kochen-Specker model, i.e. a quantum model for which ‘measurements’ correspond to (a finite set of) projectors, ‘maximal contexts’ are given by sets of mutually orthogonal projectors that resolve to the identity, and the empirical possibilities are assumed to satisfy certain functional constraints.⁵ The inconsistency between the empirical probabilities of a Kochen-Specker model and those generated by a (NC) possibilistic CHVT can then be rendered geometrically, i.e. as the impossibility of consistently ‘gluing together’ orthonormal bases (corresponding to the maximal contexts) with the following possibility assignments: each orthonormal basis has exactly one basis vector that is assigned the possibility 1 and all other basis vectors are assigned the possibility 0. For instance, an impossible gluing scenario that corresponds to the Kochen-Specker model in (Cabello et al. 1996) is depicted in Fig. 2.1, where the vertices represent possible measurement events, the rectangles represent maximal contexts, and the vertices are colored black for possibility 1, and white for possibility 0.

Such models present us with a compelling image of inconsistency which admits of an intuitive parallel with the global geometric inconsistency of ‘impossible figures’ such as Penrose’s Triangle (Penrose 1992)

³This approach goes back at least to de Finetti (1937; 1974) and can be seen more recently in Paris (2001), Kühr and Mundici (2007), Williams (2012), and Bradley (2016).

⁴We note that Spekkens (2005) generalizes the above definitions of ‘context’ and ‘noncontextual CHVT’ to apply to contexts of preparations and transformations of systems as well; however, we will only be considering measurement contexts in this paper.

⁵In particular, the assignment $h : \mathcal{P}(\mathcal{H}) \rightarrow \{0, 1\}$ of possibilities must be normative ($h(I) = 1$) and additive ($h(\bigvee_i P_i) = \sum_i h(P_i)$ for $\{P_i\}$ mutually orthogonal).

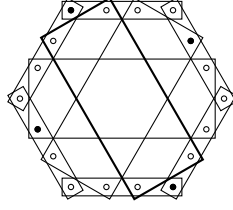


Figure 2.1: An inconsistent possibility assignment for Cabello's Kochen-Specker model.

and Escher's staircase. As we will see in Section 3.1, AB's account of Gluing Inconsistency provides both a precisification and a generalization of this core geometric intuition.

2.2 Gluing Inconsistency

Although the geometric intuition underlying Gluing Inconsistency has long been implicit in the quantum contextuality tradition, it was not until relatively recently that Abramsky and Brandenburger (**AB**) used a rigorous framework—viz. sheaf theory—to distill its essence and generalize it far beyond the confines of quantum contextuality.

The generalization offered by AB's framework is one that we shall refer to as *theory-independence*, because it allows us to consider models of a large class of theories, including *non-quantum* theories. Theory-independence can be motivated by noting that (Standard) arguments function by demonstrating an inconsistency between the 'structure of the empirical probabilities'—which happen to be generated by a quantum model—on the one hand, and a certain noncontextual CHVT, on the other. This observation suggests that one might be able to devise theory-independent versions of such arguments, in which the empirical probabilities do not necessarily come from quantum theory.

In order to do so, AB introduce a formal generalization of 'the empirical probabilities of a quantum model', which they call an 'empirical model'. Roughly speaking, an empirical model is a family of (classical) probability distributions that satisfies a generalized no-signaling condition (cf. Section 3.1 for the details). *Gluing Inconsistency* is then defined as the failure of such a family of distributions to 'glue together' to form a global/joint distribution. Furthermore, since AB show that the existence of such joint distributions is equivalent to the existence of a certain CHVT, this 'gluing failure' can be understood as a measure of the extent to which an empirical model deviates from classical probability.⁶

One of the most attractive features of AB's framework is that it clearly picks out certain features of empirical models which characterize the degree to which the model exhibits Gluing Inconsistency; henceforth, we will use the term 'contextual' to describe models with such features (we will continue to use the term 'quantum contextuality' to describe when quantum models deviate from classical probability). AB (2011) were thus able to provide the first comprehensive classification of contextuality in terms of a three-tiered hierarchy, where a higher tier (a stronger failure of gluing) implies a lower tier (a weaker failure of gluing), but not vice versa. As we will soon see, the top and middle tier of the hierarchy involve only empirical *possibilities*, whereas the lowest tier of this hierarchy is fully *probabilistic* in its characterization. In particular, the GHZ model and all Kochen-Specker models fall into the top tier, the Hardy model falls into the middle tier (but not the top), and the Bell model falls into the lowest tier (but not the middle).⁷ In sum, the 'AB hierarchy' encompasses all known quantum models that witness (Standard), and classifies the different strengths with which they deviate from classical probability.

⁶As AB note, their work thus extends arguments for quantum models along these lines that date back at least to Fine (1982).

⁷For the reader familiar with (Feintzeig 2015), we briefly note the heuristic relationship of this hierarchy with Feintzeig's hierarchy of theorems 'Kochen-Specker \Rightarrow Theorem 6 \Rightarrow Pitowsky's Theorem \Rightarrow Bell's Theorem'. The statement that 'Kochen-Specker \Rightarrow Theorem 6' concerns top-tier contextuality and additivity/convexity violations, and the statement that 'Pitowsky's Theorem \Rightarrow Bell's Theorem' concerns lowest-tier contextuality and additivity/convexity violations (we will detail these tiers, as well as the middle tier, below).

2.3 Dutch Bookability

While the notion of Gluing Inconsistency has a straightforward relationship with the quantum contextuality tradition, the same cannot be said for our second form of inconsistency, viz. Dutch Bookability. But a relationship is present nonetheless, and is most clearly seen by dividing an account of Dutch Bookability into two parts, viz. (i) the partial algebra approach; and (ii) the Dutch connection.

(i) *The partial algebra approach*

The key formal notion that Feintzeig and Fletcher (2017) use to establish a link between Dutch Books and quantum foundations is that of a Weak Probability Space (WPS), i.e. a generalization of a measure space to a setting in which the ‘set of events’ Σ is not required to have any algebraic structure (see Section 3.2 for the precise definition of a WPS) and the ‘probability function’ $\mu : \Sigma \rightarrow [0, 1]$ is a mere set function and not a measure. However, their actual use of a WPS to represent quantum models still requires *some* algebraic structure, because *co-measurable* WPS events (i.e. those that arise from measurements within the same context) are required to form an algebra. Thus, these WPS representations possess a *partial* algebraic structure in which certain algebraic relations model the co-measurability of events stemming from compatible measurements, and the absence of these algebraic relations models the non-co-measurability of events stemming from incompatible measurements—we shall call this the ‘partial algebra approach’. We stress that the WPSs with partial algebraic structure need not have anything to do with generalizations of hidden variable theories! In fact, we submit that in order to achieve complete conceptual clarity on the relation between (formal) Dutch Bookability and contextuality, it is best to simply think of a WPS as a formal device for *representing* quantum theory, akin to AB’s use of ‘empirical models’. Furthermore, only after getting clear on this relation will one be in a position to assess whether it makes sense to interpret WPSs as a kind of hidden variable theory.

The partial algebra approach has a long history within quantum foundations that is not motivated by Dutch Books, but rather by the recognition that (because of non-commuting observables) quantum models will in general give rise to measurements that fall into distinct maximal contexts; in other words, partial algebras are meant to capture the very structure that gives rise to ‘quantum contextuality’. This use of partial algebras goes at least as far back as the original Kochen-Specker paper (1975), which employs the notion of a partial Boolean algebra for this purpose. Since then, the partial algebra approach (and its associated measure theory) has been developed in various ways. One strand (Roumen 2016; Staton and Uijlen 2015; van den Berg and Heunen 2010) focuses on developing the approach in a purely algebraic manner, wherein events do not have to be realized by sets. This work has resulted in elegant (indeed, category-theoretic) constructions and powerful characterization results.

A different strand attempts to cleave very closely to the standard notion of a measure space, in which events are realized by sets. As such, it has to reckon with minutiae such as first specifying an ambient set Y that serves as the ‘sample space’ in order to specify the set of events (some subset of the power-set of Y) that carries the partial structure. One attempt to work this out in detail is provided by the notion of a ‘generalized measure space’ in (Gudder 1973). Since a WPS is a weakening of this notion, it is helpful to situate it within this strand of the partial algebra approach.

(ii) *The Dutch connection*

The connection between Dutch Books and the partial algebra approach has an antecedent in the work of Fine (1982), who can be read as suggesting that one might be able to avoid (Standard) arguments against CHVTs if one generalizes these CHVTs to account for non-co-measurability by appealing to the partial algebraic framework for probability. We will use the term *Generalized HVT* (GHVT) to refer to a particular implementation of this suggestion, viz. one that takes the relevant framework to be given by a WPS. As understood in the literature (and in order to reasonably count as a hidden variable theory) a GHVT at the very least needs to include a specification of events, probabilities of such events, and states of affairs. The first two are taken to be Σ and μ of the WPS, respectively, but very little attention has been paid to how ‘states of affairs’ are to be specified. We shall return to this point in Section 5.

The connection between WPSs and Dutch Books emerges when Feintzeig (2015) and Feintzeig and Fletcher (2017) attempt to mount a Dutch Book argument against the GHVT implementation of Fine’s suggestion. We describe this in two steps. First, they argue for the following claim:

(Dutch) There exists a quantum model Q such that any WPS that reproduces Q ’s empirical probabilities will violate a formal constraint, which in turn shows that it will be susceptible to a (formal) Dutch Book.

Regardless of whether we are discussing quantum or non-quantum models, we will refer to this type of *formal* constraint-violation as *Dutch Bookability*. We emphasize that this formal constraint-violation can be understood *without* either interpreting a WPS as a hidden variable theory or taking a subjective approach to probability, which is how we shall treat it in the rest of this paper, with the exception of Section 5.

Second, they interpret the Dutch Book normatively, thereby taking it to provide a successful no-go argument against ‘noncontextual’ GHVTs of a certain kind. It would be interesting to provide enough detail to this argument so that one can evaluate it (e.g. how should one understand ‘states of affairs’ for such a GHVT, and what does ‘noncontextual’ mean in this context?). But although we take up this subject in Section 5, our main goal is to answer a far more fundamental question concerning (Dutch), whose answer informs our assessment of the normative interpretation of Dutch Bookability.

2.4 Our motivation and results

This more fundamental question is: what is the precise relationship between (Dutch) and (Standard), as well as their theory-independent generalizations? This question has been left unanswered in the literature, although partial clues indicate that some relationship certainly exists. For instance, (Feintzeig 2015; Feintzeig and Fletcher 2017) argue by contradiction to show that the Kochen-Specker theorem implies that there exist quantum models which are formally Dutch Bookable. However, this leaves unresolved the more general question of how (Standard)—which includes non-Kochen-Specker models such as the Hardy and Bell models—is related to (Dutch), as well as the task of providing a direct and explicit account of this relationship.

A further motivation for pursuing our topic is to provide an explanation of why the particular kind of Dutch Bookability that appears in (Feintzeig 2015; Feintzeig and Fletcher 2017) stems from such a dramatic violation of additivity, in which the sample space is covered by a finite number of measure zero sets. Why should the Dutch Books of quantum theory be so extreme? In fact, the existence of a correspondence between (Dutch) and (Standard) should lead one to expect that not all such Dutch Books are so extreme, because—intuitively—not all (Standard) models are as ‘extreme’ as Kochen-Specker models in their deviation from classical probability.

Our strategy for gaining cognitive control over these issues and resolving them is to harness the power of AB’s framework for contextuality by constructing a ‘salient/physical feature-preserving’ map taking any model from AB’s framework to appropriate models in the purely formal WPS framework in which Dutch Bookability has been articulated. The use of AB’s framework is particularly illuminating in this regard because (as we will see in Section 3.1) (i) all quantum models and many non-quantum models can be represented within it; and (ii) it provides an abstract definition of contextuality and classifies contextual models into three different tiers, viz.—going from strongest to weakest—strong, logical, and probabilistic. Thus, if one succeeds in constructing an adequate map from AB’s framework to WPSs, one immediately obtains the WPS version of (i) and (ii).

Roughly speaking, our main result (see Theorem 1 and Proposition 1) can be summarized as saying that the WPS version of contextuality just is Dutch Bookability! In particular, our map sends strongly contextual empirical models (such as the Kochen-Specker model) to WPSs that maximally violate subadditivity, logically contextual empirical models (such as the Hardy model) to WPSs that violate subadditivity, and probabilistically contextual empirical models (such as the Bell model) to WPSs that violate additivity. This result then allows us to resolve the puzzle about why the Dutch Bookability that appears in (Feintzeig 2015; Feintzeig and Fletcher 2017) is so dramatic: Kochen-Specker models are not the only kind of quantum models giving rise to Dutch Books; other (Standard) models such as the Hardy and Bell models do so as well. However, Kochen-Specker models result in such a dramatic form of Dutch Bookability precisely because they witness the strongest form of contextuality. The details of these results are carefully explained in the next two sections of the paper, but the reader who is only interested in the broader conceptual story can now skip ahead to Section 5, where we discuss Feintzeig and Fletcher’s normative interpretation of (formal) Dutch Bookability as a no-go argument against GHVTs.

3 Representing empirical models

In Section 3.1, we provide an elementary account of AB’s framework for Gluing Inconsistency that is accessible to readers with basic knowledge of set theory (like AB, we will only discuss finite sets). We also use a diagrammatic visualization (see Figure 3.1) to convey why AB’s framework is so illuminating for our purposes, viz. it classifies various physical models (including quantum models) into a hierarchy of different strengths of contextuality. Section 3.2 introduces the purely formal framework of Weak Probability Spaces (WPSs) that will be used to formulate Dutch Bookability, and explains how the salient features of AB’s framework can be translated into those of a WPS. In other words, we are setting up the machinery that is required for Section 4 to demonstrate that Dutch Bookability is the WPS version of AB’s contextuality.

3.1 Gluing Inconsistency: Empirical models

The first notion that we will need to set up AB’s framework is that of an *empirical scenario*, viz. a triple (X, \mathcal{M}, O) , where X is a set of measurements, \mathcal{M} is the set of maximal contexts in the power-set $P(X)$, and O is the set of possible outcomes for each measurement (recall that a ‘context’ is a set of co-measurable/compatible measurements). We will also find it useful to refer to the set of all (possibly non-maximal) contexts, denoted \mathcal{M}' . Empirical scenarios can be represented by ‘bundle diagrams’: e.g. in Fig. 3.1, vertices of the ‘base’ represent $X = \{a, b, a', b'\}$, edges of the ‘base’ represent $\mathcal{M} = \{\{a, b\}, \{b, a'\}, \{a', b'\}, \{b', a\}\}$, and vertices over the ‘base’ represent $O = \{0, 1\}$.

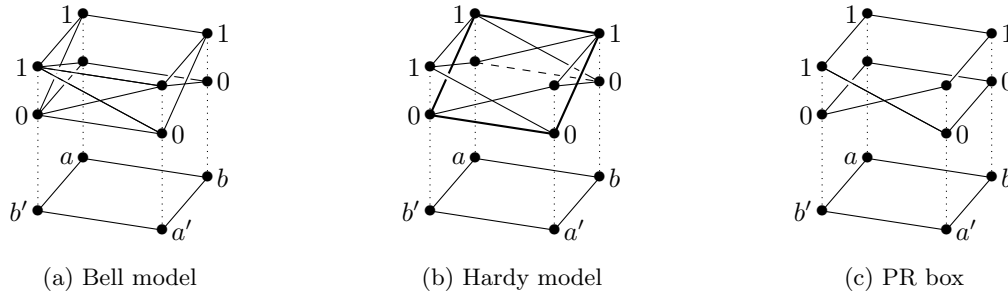


Figure 3.1: Bundle diagrams for the support of e .

Given an empirical scenario, we can define (measurement) events by means of the *event sheaf* \mathcal{E} , which maps each $U \subseteq X$ to the set of *events* $\mathcal{E}(U) := O^U$ (i.e. an event is a function from U to O), and which maps inclusions $U \subseteq U'$ to the obvious restriction $s \mapsto s|_U$, where $s \in \mathcal{E}(U')$. When we want to emphasize the geometric character of events in the AB framework, we will refer to events as *sections* over U ; if $U = X$, we say the sections are *global*, and if $U \neq X$, we say the sections are *local*. This geometric interpretation is illustrated in Fig. 3.1, where sections are represented as line segments over the base. For instance, in Fig. 3.1a, there are 4 local sections over the context $\{b', a'\}$.

We now highlight an obvious but important feature of events in the AB framework. Let $U \subseteq X$ be covered by $\{U_i\}$. We say that a family of sections $\{s_i \in \mathcal{E}(U_i)\}$ is *compatible* just in case $s_i|_{U_i \cap U_j} = s_j|_{U_i \cap U_j}$ for all i, j . The ‘sheaf property’ of \mathcal{E} amounts to the following unremarkable fact: for every compatible family $\{s_i\}$, there is a unique ‘larger’ section $s \in \mathcal{E}(U)$ that restricts to s_i on each U_i , viz. the function $s : U \rightarrow O$ that is constructed piecewise from $\{s_i\}$. In particular, when $U = X$ then the family $\{s_i\}$ gives rise to a unique global section (for instance, in Fig. 3.1b, the family of thick edges over the base glues together to form a unique global section).

Having defined events, we now proceed to define the probabilities of these events. Let a probability distribution on a finite set A be a function $d : A \rightarrow [0, 1]$ whose sum over elements of A is normalized to 1; we will use $\mathcal{D}(A)$ to denote the set of probability distributions on A . AB describe probabilities on the events of a measurement scenario by means of a ‘pre-sheaf’ \mathcal{DE} . We will only need two features of this object. First, \mathcal{DE} maps each context $U \in \mathcal{M}'$ to the set of distributions $\mathcal{DE}(U)$ on the events (sections) over U ; we will use $e_U \in \mathcal{DE}(U)$ to denote a distribution on this set. Second, let $U, U' \in \mathcal{M}'$ such that $U' \supset U$ and let A_s be the set of all $r \in \mathcal{E}(U')$ such that $r|_U = s \in \mathcal{E}(U)$. \mathcal{DE} defines a notion of ‘distribution marginalization’, viz.

$e_{U'}|_U(s) := \sum_{r \in A_s} e_{U'}(r)$ for $s \in \mathcal{E}(U)$. We say that a distribution $e_{U'} \in \mathcal{DE}(U')$ ‘restricts’ to a distribution $e_U \in \mathcal{DE}(U)$ if

$$e_U = e_{U'}|_U; \quad (3.1)$$

note that there is always some e_U for which this is the case.

We may now define the notion of an ‘empirical model’ that we earlier said would serve as a theory-independent generalization of a quantum model. Given an empirical scenario (X, \mathcal{M}, O) and the pre-sheaf \mathcal{DE} , an *empirical model* $e := \{e_C\}_{C \in \mathcal{M}}$ is a family of distributions that satisfies the following compatibility condition: for all $C, C' \in \mathcal{M}$,

$$e_C|_{C \cap C'} = e_{C'}|_{C \cap C'}. \quad (3.2)$$

AB (2011) show that, by means of a standard prescription, any quantum model induces a corresponding empirical model. For instance, the probabilities of the Bell model define a family $\{e_C\}$ for which the compatibility condition (3.2) simply amounts to the statement that model’s correlations satisfy the ‘no-signaling’ condition (for this reason, (3.2) is also referred to in the literature as a *generalized no-signaling* condition). Furthermore, AB show that the class of empirical models is much larger than just the quantum models: it includes various important *non-quantum* models which have been used as ‘foils’ in the foundations literature, such as the PR Box (Popescu 2014) and generalized Kochen-Specker scenarios like the Specker Triangle (Liang et al. 2011).

The notion of an empirical model allows us to define a three-tiered *hierarchy of contextuality* that corresponds to three different strengths of ‘gluing failure’. The first/strongest/top tier and the second/middle tier are purely ‘possibilistic’ in the sense that their definitions only use the coarse-graining of e ’s probabilities into ‘possible’ (non-zero probability) and ‘impossible’ (zero probability). We now define this sector of *possibilistic contextuality*. First, we say that the support of an empirical model e is the union of the supports $\text{supp}(e_C)$ (the set of sections that e_C assigns non-zero probabilities) for each distribution e_C . Second, we say that a global (event) section $s \in \mathcal{E}(X)$ is consistent with the support of e just in case $s|_C \in \text{supp}(e_C)$ for all C ; otherwise we say that s is inconsistent with the support of e .

- Tier 1: e is *strongly contextual* iff all its global (event) sections are inconsistent with the support of e . Well-known quantum examples of strong contextuality are provided by all (quantum) Kochen-Specker models and the GHZ model. Well-known non-quantum examples are provided by the PR Box and the Specker Triangle.
- Tier 2: e is *logically contextual* iff at least one of its global (event) sections is inconsistent with the support of e . Clearly, strong contextuality implies logical contextuality, but not vice versa (and so logical contextuality is in this sense ‘weaker’ than strong contextuality). A well-known quantum example is provided by Hardy’s model (Hardy 1993).

The geometric/topological essence of possibilistic contextuality can easily be visualized by means of ‘bundle diagrams’ of (event) sections in the support of e , as depicted in Fig. 3.1. In Fig. 3.1c we see that the PR Box is strongly contextual because no section over a maximal context can be extended to a global section of the bundle (indeed, the topological structure of this bundle is a combinatorial version of a ‘Möbius strip’, which has no global sections). Similarly, in Fig. 3.1b, we see that the Hardy model is logically contextual because there exists a section (the dotted edge) which cannot be extended to a global section of the bundle. However, the Hardy model is *not* strongly contextual, because it also contains local sections (e.g. the thick edges) that extend to global sections of the bundle.

The third and lowest tier of the hierarchy is a failure of gluing that is defined at the level of *probabilities*:

- Tier 3: $e := \{e_C\}_{C \in \mathcal{M}}$ is *probabilistically contextual* iff there does not exist a global (or joint) distribution e_X that restricts to e_C on each maximal context C .

Clearly, any empirical model that is probabilistically contextual is also probabilistically contextual—but the converse does not hold.

A well-known example of probabilistic contextuality is the Bell model, whose bundle diagram is depicted in Fig. 3.1a. Notice that all of the bundle’s sections extend to global sections, and thus it is not possibilistically contextual. But as Bell famously proved, the model is nonetheless probabilistically contextual, because the family of probability distributions (determined by Alice and Bob’s different measurement settings) does not

give rise to the joint probability distribution required for the existence of a local (NC) HVT. More generally, AB (2011) show that an empirical model e has a global distribution iff there exists a (NC) ‘factorizable’ HVT that reproduces the probabilities of e : this result thus constitutes a theory-independent generalization of (Standard).⁸

In sum, the hierarchy of contextuality is:

$$\text{Strong} \implies \text{Logical} \implies \text{Probabilistic} \quad (3.3)$$

In the next subsection, we explain how the structural features that allow us to define this hierarchy can be represented within the WPS framework.

3.2 Weak Probability Spaces: Representations of empirical models

Feintzeig and Fletcher (2017) assume that Generalized HVTs (GHVTs) will at least have the structure of a Weak Probability Space (WPS). Thus, in order to demonstrate that GHVTs are Dutch Bookable, they provide a method of representing a quantum Kochen-Specker model within the WPS framework. The main task of this section will be to obtain a theory-independent generalization of this method by constructing a WPS representation of an empirical model.

Recall that a probability space is a triple (Y, Σ, μ) where Σ is an algebra of events (viz. a σ -field on Y), and $\mu : \Sigma \rightarrow [0, 1]$ is a probability measure.⁹ A WPS is also a triple (Y, Σ, μ) , but it generalizes a measure space by replacing the algebra of events with a mere set of events $\Sigma \subseteq P(Y)$ and by replacing the measure with a mere set function $\mu : \Sigma \rightarrow R \supseteq [0, 1]$; we will nonetheless find it convenient to refer to this set function as a ‘measure’ when there is no danger of confusion. While this definition of Σ is incredibly weak, remember that it only functions as a template: when a WPS is used to represent an empirical model, this template will need to be filled in with the relevant co-measurability relations.

We now proceed to explain how events in AB’s framework (‘AB-events’) can be *represented* in a WPS, and how the WPS events (‘ Σ -events’) can then be constructed from a proper subset of these representations. For brevity, we will use ‘ s_U ’ to denote an AB-event section in $\mathcal{E}(U)$.

Let e be an empirical model. An *event representation* of e , denoted \mathcal{E}^\dagger , consists of a sample space Y and an injective ‘event transfer map’

$$\bar{E} : \bigcup_{U \subseteq X} \mathcal{E}(U) \rightarrow P(Y). \quad (3.4)$$

that satisfies the following analog of the ‘sheaf’ property for AB-events: for any $U \subseteq U' \subseteq X$,

$$\bar{E}(s_{U'}) = \bigcap_{\{s_U : s_{U'}|_U = s_U\}} \bar{E}(s_U) \neq \emptyset. \quad (3.5)$$

Recall that $s_{U'}$ is uniquely determined by the AB-events $\{s_U\}$ that it restricts to; (3.5) simply encodes this set-theoretically by representing $s_{U'}$ as the conjunction of $\{\bar{E}(s_U)\}$ (the intersections are required to be non-trivial because any of these events may be possible).

At this juncture, one could impose two further conditions to ensure that an event representation \mathcal{E}^\dagger fully encodes the combinatorial structure of the event sheaf \mathcal{E} of e , viz.

Strong Mutual Exclusivity For any distinct $s, s' \in \mathcal{E}(U)$ and any $U \subseteq X$, $\bar{E}(s) \cap \bar{E}(s') = \emptyset$.

Exhaustiveness For any $U \subseteq X$, $\bigcup_{s \in \mathcal{E}(U)} \bar{E}(s) = Y$.

Call an event representation satisfying these additional conditions a *combinatorial event representation*.¹⁰ In what follows, we will prove our main result (Theorem 1) about the relationship between contextuality and

⁸AB define ‘factorizability’ as follows: consider the HVT specified by some distribution $h_\Lambda \in \mathcal{D}_R(\Lambda)$ over ontic states $\lambda \in \Lambda$, where ontic states induce response functions $h_C^\lambda \in \mathcal{D}_R \mathcal{E}(C)$ such that $e_C = \sum_{\lambda \in \Lambda} h_C^\lambda \cdot h_\Lambda(\lambda)$ for all $C \in \mathcal{M}$. This HVT is *factorizable* (Abramsky and Brandenburger 2011, Section 8) iff, for any $\lambda \in \Lambda$ and for any $s \in \mathcal{E}(C)$ given any $C \in \mathcal{M}$, $h_C^\lambda(s) = \prod_{m \in C} h_C^\lambda|_{\{m\}}(s|_{\{m\}})$.

⁹Recall that a (probability) measure is an additive function $\mu : \Sigma \rightarrow \mathbb{R}_{\geq 0}$ such that $\mu(\emptyset) = 0$ and $\mu(Y) = 1$. A σ -field Σ on Y is a collection of subsets of Y containing \emptyset that is closed under complementation and countable unions and intersections.

¹⁰Every empirical model has a minimal combinatorial event representation; we construct one example in the Appendix.

Dutch Bookability without assuming that the event representation is combinatorial (this degree of generality allows us to keep our treatment as close to that of Feintzeig and Fletcher as possible). However, we also show (Props. 1 and 2) that if one assumes the combinatorial structure, one can give a simpler argument for a stronger result, which has converses.

In order to avoid the clutter of repeatedly writing \bar{E} , we will introduce the notation $S_U := \bar{E}(s_U)$ for event representations, as well as $\mathcal{E}^\dagger(U) := \{\bar{E}(s_U) : s_U \in \mathcal{E}(U)\}$ for the set of event representations over $U \subseteq X$.

We note that there is a ‘duality’ between AB-events and their WPS representations: if $U \subset U'$ and $s_U = s_{U'}|_U$, then by (3.5), $S_U \supset S_{U'}$. Furthermore, just as the event sheaf \mathcal{E} comes with a notion of restriction, \mathcal{E}^\dagger possesses a dual notion of extension, i.e. $S_{U'}|^{U'} = S_U$, where $s_U = s_{U'}|_U$. And just as every compatible family of AB-event sections $\{s_i \in \mathcal{E}(U_i)\}$ (where $\{U_i\}$ covers X) gives rise to a unique global section s_X , any compatible family $\{S_i \in \mathcal{E}^\dagger(U_i)\}$ (i.e. a family that satisfies $S_i|^{U_i \cap U_j} = S_j|^{U_i \cap U_j}$ for all i, j) has a unique intersection S_X such that $S_X|^{U_i} = S_i$ for all i .

To prepare ourselves for the definition of a WPS representation of e , let us consider how co-measurability relations can be faithfully transferred from the AB-events to the set of Σ -events that we now wish to construct. First, recall that in the AB framework, ‘co-measurable events’ are sections over sets of measurements that are jointly contained in a context, and ‘non-co-measurable events’ are sections over sets of measurements that are not. Thus, in order to transfer the structure of co-measurable events from the AB framework to the WPS framework, we will restrict the domain of \bar{E} to AB-events over contexts, thereby defining the map $E : \bigcup_{U \in \mathcal{M}'} \mathcal{E}(U) \rightarrow P(Y)$. It is this map that we will now use to construct the event set Σ of a WPS representation.

For any context $U \in \mathcal{M}'$, let Σ_U denote the algebra generated by the collection of sets $\bigcup_{x \in U} \mathcal{E}^\dagger(x)$.¹¹ The set of WPS events Σ is then defined as the union $\bigcup_{U \in \mathcal{M}'} \Sigma_U$, where Σ_U is treated as a collection of sets. Non-co-measurable Σ -events are thus elements of Σ which are not jointly contained in some Σ_U .

A *WPS representation* of e , denoted e^\dagger , is a triple $(\mathcal{E}^\dagger, \Sigma, \mu)$ such that (Y, Σ, μ) form a WPS that satisfies the following three conditions. First, we require that μ allows us to define a probability space for each algebra Σ_U :

Weak Classicality (WC) Let $e_U^\dagger := \mu|_{\Sigma_U}$. For any context $U \in \mathcal{M}'$, $(Y, \Sigma_U, e_U^\dagger)$ is a probability space.

Second, in order to ensure that e^\dagger accurately encodes the empirical probabilities of e , we require:

Empirical Consistency (EC) For any context $U \subset C \in \mathcal{M}$ (where C is a maximal context) and any AB-event $s \in \mathcal{E}(U)$,

$$e_U^\dagger(S) = e_C|_U(s). \quad (3.6)$$

Third, since e treats events over contexts as mutually exclusive, we require:

Mutual Exclusivity (ME) For any distinct AB-events $s, s' \in \mathcal{E}(U)$, $e_U^\dagger(S \cap S') = 0$.

This concludes our construction of a WPS representation e^\dagger of an empirical model e .¹²

As one might expect, any WPS representation e^\dagger will obey duals of the marginalization condition (3.1) and the compatibility condition (3.2) for e . To describe the dual of (3.1), let $U \subseteq U' \subseteq C$ and $s \in \mathcal{E}(U)$, let A_S be the set of $R \in \mathcal{E}^\dagger(U')$ such that $R|_U = S$, and define ‘marginalization for extension’ as $e_{U'}^\dagger|_U(S) := \sum_{R \in A_S} e_{U'}^\dagger(R)$. By (EC), it immediately follows that

$$e_{U'}^\dagger|_U = e_U^\dagger, \quad (3.7)$$

which is the dual of (3.1). The dual of (3.2) is simply the statement that WPS representations satisfy the following compatibility condition: for all maximal contexts $C, C' \in \mathcal{M}$,

$$e_C^\dagger|^{C \cap C'} = e_{C'}^\dagger|^{C \cap C'}. \quad (3.8)$$

¹¹Here and henceforth, we will use x to denote the singleton $\{x\}$ when this meaning is clear from context.

¹²For the precise relationship between Feintzeig and Fletcher’s *weak hidden variable representations* and our *WPS representations*, see the Appendix.

We now have two ways of formalizing the empirical probabilities of a wide class of physical models: AB’s sheaf-theoretic approach, on the one hand, and the above ‘dual’ WPS approach, on the other. In the next section, we will argue that AB’s hierarchy of contextuality (or Gluing Inconsistency) corresponds to a hierarchy of Dutch Bookability in the WPS approach.

4 The hierarchies

We begin this section by situating our result with respect to previous work on hierarchies of constraint-violations leading to Dutch Books, albeit work that has only been done for *abstract* WPSs, and not WPS representations of *physical* models. Recall that Feintzeig and Fletcher (2017) show that any WPS representation of a Kochen-Specker (i.e. a quantum) model will violate a constraint (in their terminology: the ‘no finite null cover’ (NFNC) condition), which amounts to a maximal violation of ‘subadditivity’ in a sense that we will make precise below. This of course raises the question of why this violation is maximal, and whether there might be a hierarchy of different violations enjoyed by quantum and other physical models.

Feintzeig (2015) and Wroński and Godziszewski (2017) have attempted to place some constraints on this question by reasoning abstractly about WPSs. Feintzeig (2015) notes that even if an abstract WPS does not maximally violate subadditivity, it could still exhibit a weaker violation of subadditivity that gives rise to a formal Dutch Book. This leads him to conjecture that an abstract WPS satisfies subadditivity iff it is not Dutch Bookable—if this were true, a WPS representation of a physical model might both offer a non-classical alternative to HVTs and avoid Dutch Books. However, Wroński and Godziszewski (2017) show that this conjecture is false by constructing an abstract WPS that is both Dutch Bookable and subadditive. More importantly, they show that an abstract WPS is Dutch Bookable just in case it can be modeled by classical probability (cf. Section 4.1 below).

We concur with Wroński and Godziszewski (2017) that this exchange demonstrates the futility of simultaneously using WPSs as a framework for GHVTs and stipulating that probabilities in these theories must avoid formal Dutch Books. However, for all this talk of an abstract hierarchy, it could still be the case that no interesting class of physical models gives rise to Dutch Books except by maximally violating subadditivity! In other words, it might be the case that as regards WPS representations of *physical models*, no discussion of a hierarchy is necessary beyond the original Dutch Book violation demonstrated by Feintzeig and Fletcher (2017).

In what follows, we will show that this is not the case by arguing for a correspondence between AB’s hierarchy of contextuality and a hierarchy of ‘formal constraint’-violations that lead to Dutch Bookability. Since Feintzeig (2015) emphasizes formal constraints with a measure-theoretic flavor, viz. additivity conditions, and the approach of Wroński and Godziszewski (2017) suggests convexity constraints, we will discuss both aspects in Sections 4.1 and 4.2 respectively.

4.1 Contextuality and additivity-violation

We now introduce the relevant additivity conditions that will be violated in our hierarchy. These violations are always witnessed by the ‘defect of subadditivity’ for some (finite) set V of Σ -events,

$$\mathfrak{a}(V) := \mu\left(\bigcup V\right) - \sum_{a \in V} \mu(a). \quad (4.1)$$

A WPS violates subadditivity just in case there is a V such that $\mathfrak{a}(V) > 0$, and it maximally violates subadditivity just in case there is a V such that $\mathfrak{a}(V) = 1$. Note that a ‘finite null cover’ of a WPS representation—a finite collection of measure-zero sets covering the sample space Y —provides one example of a maximal violation of subadditivity. A WPS violates additivity just in case there is some (finite) set V of *disjoint* Σ -events such that $\mathfrak{a}(V) \neq 0$.

In order to construct a hierarchy of additivity-violation, we will also need the notion of an ‘extension’ of a WPS $W \equiv (Y, \Sigma, \mu)$, as well as the more refined notion of a monotonic extension. An *extension* of W is a WPS (Y, Σ', μ') , where Σ' contains the algebra generated by Σ , and $\mu'|_{\Sigma} = \mu$.¹³ A *monotonic extension*

¹³A classical extension is just an extension where μ' is a probability measure.

of W is an extension (Y, Σ', μ') of W for which $\mu' : \Sigma' \rightarrow [0, 1]$ is a monotonic function, i.e. if $A \subseteq B$, then $\mu'(A) \leq \mu'(B)$.¹⁴

In the case of *abstract* WPSs, it is easy to construct a hierarchy of additivity-violation that is tantalizingly similar to AB's hierarchy of contextuality. It is clear that a maximal violation of subadditivity implies a violation of subadditivity, which in turn implies that any of the monotonic extensions of the WPS in question must violate additivity. Given this observation, one may naturally expect the following hierarchy for WPS representations of empirical models: strong contextuality will correspond to a maximal violation of subadditivity; logical contextuality will correspond to a violation of subadditivity; and probabilistic contextuality will correspond to a violation of additivity for monotonic extensions of e^\dagger . The following theorem shows that this expectation is realized.

Theorem 1. *Let e be an empirical model and let e^\dagger be any of its WPS representations.*

1. *If e is strongly contextual, then e^\dagger maximally violates subadditivity.*
2. *If e is logically contextual, then e^\dagger violates subadditivity.*
3. *If e is probabilistically contextual, then any monotonic extension of e^\dagger violates additivity.*

In fact, if one makes the stronger assumption that e^\dagger is *combinatorial*, meaning that its event representation is combinatorial (cf. Section 3.2), then one obtains an *equivalence* between each tier of contextuality and a corresponding condition that witnesses additivity-violation. We now state these conditions and the equivalence result.

Given that strong and logical contextuality only concern the possibilistic data of an empirical model e , we expect the corresponding (sub)additivity-violations to depend only on the possibilistic data of e^\dagger . To characterize such violations, we will introduce the notion of an *additive cover*, viz. a finite set $A \subseteq \Sigma$ of mutually disjoint Σ -events that is additive (so $\mathbf{a}(A) = 0$) and that covers the sample space Y . Now say e^\dagger *strongly \mathcal{V} -violates subadditivity* just in case a finite set V of measure-zero elements in $\mathcal{V} \subseteq \Sigma$ covers *all* the non-measure-zero elements of an additive cover $A \subseteq \mathcal{V}$ (and so $\mathbf{a}(V) = 1$). Additionally, say e^\dagger *logically \mathcal{V} -violates subadditivity* just in case a finite set V of measure-zero elements in $\mathcal{V} \subseteq \Sigma$ covers *some* non-measure-zero element a of an additive cover $A \subseteq \mathcal{V}$ (and so $\mathbf{a}(V \cup A \setminus \{a\}) > 0$).

Finally, in order to characterize additivity-violations stemming from the probabilistic data of e^\dagger , we turn to monotonic extensions. Say e^\dagger *\mathcal{V} -violates additivity* just in case the algebra generated by \mathcal{V} contains a collection V of disjoint sets that violates additivity in any monotonic extension of e^\dagger (and so $\mathbf{a}(V) \neq 0$ in these extensions).

In the following proposition, the above conditions are witnessed by the set $\mathcal{V}_\mathcal{M}$ of Σ -events corresponding to AB-events over maximal contexts (i.e. $\mathcal{V}_\mathcal{M} := \{S : S \in \mathcal{E}^\dagger(C), C \in \mathcal{M}\}$).

Proposition 1. *Let e be an empirical model and let e^\dagger be any of its combinatorial WPS representations.*

1. *e is strongly contextual iff e^\dagger strongly $\mathcal{V}_\mathcal{M}$ -violates subadditivity.*
2. *e is logically contextual iff e^\dagger logically $\mathcal{V}_\mathcal{M}$ -violates subadditivity.*
3. *e is probabilistically contextual iff e^\dagger $\mathcal{V}_\mathcal{M}$ -violates additivity.*

Note that these violations of additivity are *invariant*: if *one* combinatorial WPS representation of e (strongly or logically) $\mathcal{V}_\mathcal{M}$ -violates (sub)additivity, then so do *all* combinatorial WPS representations of e . There is no apparent reason that this need be the case for Theorem 1, where we consider *arbitrary* WPS representations (and one may well consider this to be another sort of pathology of these structures).

The proof of Theorem 1 involves some set-theoretic yoga. We relegate the details to the Appendix, but it is possible to briefly convey an intuitive idea of how the argument works. First, observe that Y contains two measure-zero collections of Σ -events which do not represent any AB-events, viz. the collection of ‘contradictory events’ D_1 and the collection of ‘non-existent outcomes’ D_2 (cf. Eqns (5.1) and (5.2) of the Appendix for the full definitions). In order to translate the definitions of strong/logical/probabilistic contextuality into a WPS setting, we thus excise $\bigcup(D_1 \cup D_2)$ from Y to obtain Z , which represents all the

¹⁴Note that μ' is really a ‘fuzzy measure’ in the sense of (Murofushi and Sugeno 1989).

properties that we care about.¹⁵ The translation is then carried out by relying on a key lemma (Lemma 1), which tells us that any $z \in Z$ is contained in some S_X (representing a global section s_X). For instance, this lemma lets us translate ‘logical contextuality’ into the statement that some subset of Z is contained in a collection of measure zero sets, which lets us construct the relevant sub-additivity-violating collection.

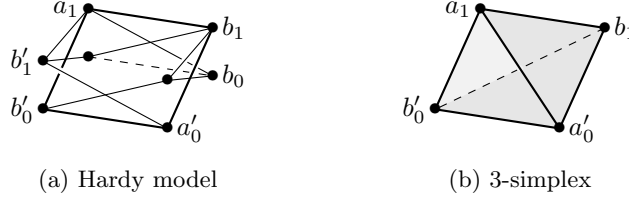


Figure 4.1

In the case where e is possibilistically contextual, there is also a beautiful topological interpretation of e^\dagger 's violation of subadditivity that mirrors the bundle diagrams of Fig. 3.1. For instance, consider e^\dagger for the Hardy model. Fig. 4.1a is a ‘nerve diagram’ of the Σ -events in the support of μ , meaning that n -simplices represent n -intersections of Σ -events. Notice that diagram is essentially the bundle diagram 3.1b, except that the AB-event $a \mapsto 0$ has been replaced by the Σ -event (or 0-simplex) $a_0 := E(a \mapsto 0)$, the AB-event $\{a \mapsto 0, b \mapsto 0\}$ with the Σ -event (or 1-simplex) $a_0 \cap b_0$, and so on. The relevant violation of subadditivity can then be understood as stemming from the fact that the dashed 1-simplex is not the edge of a 3-simplex (the WPS-analogue of a global section) whose edges are in the support of μ .¹⁶

Two important corollaries immediately follow from Thm. 1. First, since AB (2011) have shown that the Hardy model is logically contextual but not strongly contextual, and that the Bell model is probabilistically contextual but not logically contextual, our result shows that there are quantum models which ‘inhabit’ lower tiers of the hierarchy of additivity-violation. It also extends the original NFNC-violation result of Feintzeig and Fletcher (2017) by showing that *non-quantum* models like the PR Box have WPS representations which maximally violate subadditivity.

To state the second corollary, we will need a final notion of ‘extension’ discussed in (Feintzeig 2015; Wroński and Godziszewski 2017): a *classical extension* of W is an extension (Y, Σ', μ') that is a probability space, i.e. Σ' is an algebra, and μ' is additive. The corollary is that e is contextual (either strongly, or logically, or probabilistically) only if its WPS representations do not have a classical extension. Thm. 2 of (Wroński and Godziszewski 2017)—a minor adaptation of a result of Paris (2001)—states that an abstract WPS is Dutch Bookable iff it does not have a classical extension. Thus, if e is contextual then e^\dagger is Dutch Bookable. The complete set of relations between our two hierarchies is shown in Fig. 4.2.

We note that the aforementioned result of Paris (2001) is essentially one about ‘convexity’, which lies at the heart of the formal concept of a Dutch Book. The next section reviews formal Dutch Books and describes the hierarchy of contextuality in terms of a hierarchy of convexity-violation.

4.2 Dutch Bookability and convexity-violation

The violation of the formal ‘No Dutch Books’ constraint and its standard normative interpretation is well-trodden territory, so we shall be brief in our summary. Our main aim is to stress that since formal Dutch Bookability is a statement about convexity-violation, it follows that AB’s hierarchy of contextuality can be straightforwardly understood as a hierarchy of convexity-violation (cf. Prop. 2 below). In order to separate this point from normative issues, we will defer our review of the standard normative interpretation of formal Dutch Books to Section 5.

Let (Y, Σ, μ) be a WPS. To succinctly state the standard definition of a formal Dutch Book (adapted to WPSs), it will be useful to define the (finite) set of functions $\mathbb{V} := \{ V_y = \chi_{(\cdot)}(y) : \Sigma \rightarrow \{0, 1\} \}$, where $y \in Y$ and $\chi_A : Y \rightarrow \{0, 1\}$ is the characteristic function of $A \in \Sigma$. Each function $V_y \in \mathbb{V}$ can be identified

¹⁵This rough strategy of ‘excising sets’ was applied by Feintzeig and Fletcher (2017) only to Kochen-Specker models. We go much further, generalizing the strategy using sheaf theory so that it applies to *all* forms of contextuality.

¹⁶Clearly, this suggests a cohomological formulation of additivity-violation, although we have not the room to treat this aspect here.

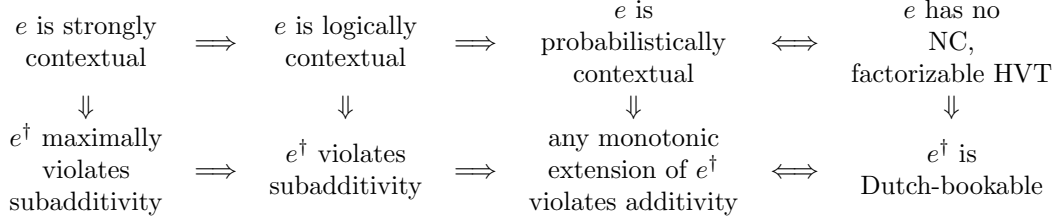


Figure 4.2: The hierarchy of contextuality (top) and the hierarchy of Dutch Bookability from additivity-violation (bottom).

with the ‘elementary/atomic event’ specified by an element y of the sample space Y . We say that a WPS violates a ‘No Dutch Books’ constraint (or: is Dutch Bookable) iff there exists a function $s : \Sigma \rightarrow \mathbb{R}$ and a finite collection of sets $\mathcal{V} \subset \Sigma$ such that

$$\text{for any } V_y \in \mathbb{V}, \sum_{A \in \mathcal{V}} s(A) \cdot (V_y(A) - \mu(A)) < 0. \quad (4.2)$$

It is also common to leave (Y, Σ) implicit and simply say that the set-function μ is Dutch Bookable.

It is a well-known fact that condition (4.2) is a statement about convexity-violation. More precisely: let $\text{Conv}(\mathbb{V})$ be the convex hull of \mathbb{V} , i.e. the set of all convex combinations of functions in \mathbb{V} . We will say that μ is \mathbb{V} -convex iff $\mu \in \text{Conv}(\mathbb{V})$. Then, as originally noted by de Finetti (1974), (for finite \mathbb{V}) μ is not Dutch Bookable iff μ is \mathbb{V} -convex.¹⁷

With this background in place, it is easy to see that AB’s hierarchy of contextuality can be interpreted as a hierarchy of obstructions to \mathbb{V} -convexity. For $\mathcal{V} \subseteq \Sigma$, we say that a WPS representation e^\dagger \mathcal{V} -violates \mathbb{V} -convexity so long as $\mu|_{\mathcal{V}}$ cannot be written as a convex sum of $V_y|_{\mathcal{V}}$ for $V_y \in \mathbb{V}$; this will constitute the lowest tier of the hierarchy. In order to describe the higher tiers, let us introduce the Boolean addition operator \oplus , i.e. $1 \oplus 1 = 1$, $0 \oplus 1 = 1 \oplus 0 = 1$, and $0 \oplus 0 = 0$ (logical sums of functions are conducted pointwise). Let χ_μ be the characteristic function of the support of μ (i.e. for $A \in \Sigma$, $\chi_\mu(A) = 0$ if $\mu(A) = 0$ and $\chi_\mu(A) = 1$ otherwise). We will say that e^\dagger logically \mathcal{V} -violates \mathbb{V} -convexity iff $\chi_\mu|_{\mathcal{V}}$ is not a logical sum of $V_y|_{\mathcal{V}}$ for $V_y \in \mathbb{V}$; clearly, in such a case, no convex combination of V_y will yield μ . In this case, however, it may still be that

$$\chi_\mu|_{\mathcal{V}} = f + \bigoplus_{V_y \in \mathbb{V}'} V_y|_{\mathcal{V}} \quad (4.3)$$

for some non-empty $\mathbb{V}' \subseteq \mathbb{V}$ and some function $f : \mathcal{V} \rightarrow \{0, 1\}$. If (4.3) cannot be satisfied for \mathcal{V} , then we will say that e^\dagger strongly \mathcal{V} -violates \mathbb{V} -convexity. Given these definitions, one immediately obtains the following correspondence between contextuality and convexity-violation (we sketch a proof in the Appendix):

Proposition 2. *For any empirical model e and any of its combinatorial WPS representations e^\dagger ,*

1. *e is strongly contextual iff e^\dagger strongly $\mathcal{V}_{\mathcal{M}}$ -violates \mathbb{V} -convexity.*
2. *e is logically contextual iff e^\dagger logically $\mathcal{V}_{\mathcal{M}}$ -violates \mathbb{V} -convexity.*
3. *e is probabilistically contextual iff e^\dagger $\mathcal{V}_{\mathcal{M}}$ -violates \mathbb{V} -convexity.*

By de Finetti’s result, the probability measure μ of e^\dagger violates \mathbb{V} -convexity iff it is Dutch Bookable; thus, this hierarchy of violations of \mathbb{V} -convexity is a hierarchy of conditions witnessing Dutch Bookability.

5 Is Dutch Bookability normative?

In this paper, we have demonstrated a correspondence between AB’s hierarchy of contextuality and a hierarchy of Dutch Bookability for WPS representations. This shows that (Standard) arguments against CHVTs

¹⁷See also Theorem 2 of Paris (2001).

can be mapped to corresponding (Dutch) arguments for WPS representations. In this final section, we turn to a partial evaluation of whether formal Dutch Bookability can be used to provide a normative argument against GHVTs.

Feintzeig and Fletcher (2017) seek to use (Dutch) arguments to demonstrate that the probabilities of hidden variable theories based on WPSs—which we have labeled GHVTs—cannot be given a ‘rational’ subjective interpretation. To unpack this claim, first recall that the existence of a ‘finite null cover’ constitutes a maximal violation of subadditivity, the strongest rung of the Dutch Bookability hierarchy (pictured in Figure 4.2). Now, note that Feintzeig and Fletcher’s specific claim is that ‘whenever a weak probability space [has a finite null cover], anyone who sets their beliefs according to that probability space will be irrational in the sense that she will accept as fair a series of bets on which she is guaranteed to lose money’ (2017, p. 308).

This is a dramatic claim! As Feintzeig and Fletcher note, WPSs are widely used in the literature, and several authors explicitly use such structures as the basis for hidden variable theories (Craig et al. 2006; Feintzeig and Fletcher 2017; Suppes and Zanotti 1991). Is a subjective approach to the probabilities in any such GHVT irrational, as Feintzeig and Fletcher seem to suggest? We will not provide a complete assessment of this claim here; however, we wish to highlight that on one natural and plausible reconstruction, the soundness of Feintzeig and Fletcher’s normative argument requires *at least* a very specific approach to the ‘hidden variable’ interpretation of a WPS.

To unpack this approach, it is helpful to consider what GHVTs (and their classical counterparts) *could* be. As standardly construed in the literature, and by our lights, a hidden variable theory (e.g. a CHVT or a GHVT) should include at least the following schematic ingredients:

- (i) A structure of *events*;
- (ii) *Probabilities* of these events;
- (iii) *States of affairs* and their associated *response functions* (recall that a response function specifies the probability of an event conditional on the system’s being in some particular state);
- (iv) A *relationship* between the probabilities of events and the response functions.

For instance, Abramsky and Brandenburger flesh out this schema for a CHVT as follows. Let $\mathcal{E} = \{E\}$ be the set of events described in Section 3.1 (the reader re-joining us from Section 2 may safely view \mathcal{E} as a set without any additional structure)—this specifies (i). Let the Born rule specify the empirical probabilities $p(E|C)$, where $C \in \mathcal{M}$ is a maximal set of compatible measurements that can yield E —this specifies (ii). Now specify (iii) by introducing a (finite) set of states of affairs Λ . We take each $\lambda \in \Lambda$ to specify a response function $p(\cdot | \lambda, C)$ for each $C \in \mathcal{M}$, which maps all events that can be measured in the specified context to weights in the unit interval. Finally, introduce functions $p(\cdot) : \Lambda \rightarrow [0, 1]$ such that $\sum_{\lambda} p(\lambda) = 1$, which describe the probabilities that the preparation of the system yields a given state of affairs. (iv) is then specified by requiring that the empirical probabilities are reproduced by means of the following functional relationship:

$$p(E|C) = \sum_{\lambda \in \Lambda} p(\lambda)p(E|\lambda, C). \quad (5.1)$$

This approach allows for a transparent definition of ‘noncontextual’, which is also a standard definition in the quantum foundations literature. The definition is that response functions are noncontextual just in case $p(E|\lambda, C) = p(E|\lambda, C')$ for all contexts $C, C' \in \mathcal{M}$ that can yield the measurement event E .¹⁸

How might GHVTs satisfy the above schema? At the very least, all the WPS-users mentioned above use the set Σ of WPS events to represent the structure of measurement events, and they use the WPS ‘measure’ μ to represent the empirical probabilities given by the Born rule (the usual probabilistic rule of quantum theory). This specifies (i) and (ii), and we take these choices to be uncontroversial.

However, Feintzeig and Fletcher do not specify (iii) and (iv) (at least explicitly). We now provide a charitable reconstruction of how one might do so in order to construct a valid (Dutch) argument against the ‘coherence’ of GHVTs. Let us first recall the standard definition of Dutch Bookability generalized to WPSs

¹⁸Note that this definition of ‘noncontextual’ aligns with Spekkens’s (2005) notion of ‘measurement noncontextuality’ (albeit without his operational definition of ‘measurement contexts’).

(c.f. Section 4.2) for the reader re-joining us from Section 2. Let Y be the sample space of a WPS, let Σ (a subset of the power set of Y) be the set of WPS-events, and let \mathcal{V} be some finite subset of Σ . Now, in the spirit of classical probability theory, define for each point y in the sample space Y an ‘elementary event’ $V_y := \chi_{(\cdot)}(y) : \Sigma \rightarrow [0, 1]$, where $\chi_{(A)}(y)$ is the characteristic function of the WPS-event A evaluated at y . Now, we say a WPS set-function $\mu : \Sigma \rightarrow [0, 1]$ violates a ‘No Dutch Books’ constraint (or: is Dutch Bookable) just in case there is some set \mathcal{V} and some function $s : \mathcal{V} \rightarrow \mathbb{R}$ specifying real-valued ‘stakes’ such that

$$\text{for any } V_y \in \mathbb{V}, \sum_{A \in \mathcal{V}} s(A) \cdot (V_y(A) - \mu(A)) < 0. \quad (5.2)$$

There is a standard ‘normative Dutch Book’ story from formal epistemology whose aim is establish that Dutch Bookability is pathological (de Finetti 1974; Hájek 2008). This story interprets equation (5.2) as a betting game. Suppose a bookie proposes a game wherein each WPS event $A \in \mathcal{V}$ is valued at $s(A)$ dollars. If you agree to the game, then if V_y describes the state of the world, you will receive (or be forced to pay the bookie, if $s(A)$ is negative) $s(A) \cdot V_y(A)$ dollars when the game ends. However, to play the game, you must put down (or agree to be paid by the bookie, if $s(A)$ is negative) $s(A) \cdot \mu(A)$ dollars. If $\mu(\cdot)$ truly reflects your expectations for the values of $V_y(\cdot)$ (the thought goes), then you should accept this game as ‘fair’. But then if equation (5.2) holds, you will accept as fair a series of bets on which you will lose money no matter what the world turns out to be like.

There are many reasonable philosophical worries about this story that we will not address here.¹⁹ Instead, our concern is with the following point: in order for normative Dutch Books to make any sense for WPSs, one has to take seriously the idea that the elementary events of a WPS have something interesting to say about *what the world is like*. What could this something be? By our lights, the most reasonable answer is to take the elementary events of a WPS to be response functions, thereby specifying (iii). This move also gives a sense to Feintzeig and Fletcher’s assertion that the relevant GHVTs are ‘non-contextual’ (a term they do not define): these response functions are non-contextual in the sense defined above (just after (5.1)).

The above specification of (iii) allows us to clearly lay out a certain subjective approach to specifying (iv), and to explain how a (Dutch) argument can be launched against it. The approach interprets the ‘empirical probabilities’ μ as an agent’s subjective expectations for the values of response functions. (iv) is then specified by an argument that the functions μ are all and only those expectations that satisfy a certain coherence condition, which is standardly taken to be the avoidance of Dutch Books (cf. de Finetti (1974)). In other words, a GHVT whose response functions are taken to be $V_y(\cdot)$ has a ‘coherent’ μ just in case constraint (5.2) is satisfied. On this basis, we can reconstruct Feintzeig and Fletcher’s original (under-specified) argument as saying that, given (i)–(iv), (Dutch) shows that the relevant GHVT is ‘incoherent’.

We now emphasize two morals concerning this reconstruction of the normative (Dutch) argument against GHVTs. First, Feintzeig and Fletcher (2017) motivate GHVTs by suggesting (*à la* Pitowsky 1989) that they can escape one (Standard) argument—Bell’s theorem—by weakening the strictures of classical probability theory. But our reconstruction requires specifying NC, OD response functions for GHVTs. Any HVT with such response functions is ‘factorizable’—very roughly, joint distributions for several outcomes in one context factor as the product of the probabilities of each outcome. And by Fine’s theorem (1982), if quantum probabilities witness a violation of Bell inequalities, then there *cannot be* such an HVT—so long as the quantum probabilities ought to be reproduced by equation (5.1).²⁰ It was noted in Section 3.1 that the non-existence of such an HVT is an example of the third and lowest tier of AB’s hierarchy of contextuality. The more general moral, then, is that GHVTs with NC, OD response functions are susceptible to (Standard)—and *thereby* (Dutch)—arguments; thus, (Dutch) arguments simply re-express (Standard) arguments within the language of the WPS framework.

Second, while Feintzeig and Fletcher seem to suggest that a subjective approach to the probabilities of *any* GHVT (representing a Kochen-Specker model) is irrational because it is Dutch Bookable, this consequence can be avoided by any WPS-user who divests elementary events of physical meaning and specifies response functions in some other way. For example: if response functions are given by wavefunctions in a finite-dimensional Hilbert space, then all and only the probabilities given by density matrices avoid Dutch Books.²¹

¹⁹For some of these, see Kaplan (1998).

²⁰See Kunjwal (2015) for a lucid discussion of this point in the ontological models framework.

²¹See Steeger (2018) for a demonstration of this point, as well as a generalization of it, in the C^* -algebraic framework.

We conclude by noting that once enough of the mathematical features of a WPS (such as ‘elementary events’) are stripped of physical meaning, one may well begin to question the usefulness and interest of WPSs as the backbone of a hidden variable theory. Our own view is that WPSs are primarily of interest because they are a quasi-probabilistic framework within which we can represent quantum (and non-quantum) models and classify different strengths of contextuality, analogous to Abramsky and Brandenburger’s sheaf-theoretic framework—this is the sense in which our results about the hierarchy of Dutch Bookability are of foundational significance. But such structures can be interesting while still failing to be useful for a hidden variables approach to quantum theory.

Acknowledgements

We thank Samson Abramsky, Jeremy Butterfield, Bryan Roberts, Alex Pruss, and Alan Hájek for helpful discussions of this material. Nicholas Teh thanks audiences at the Rutgers Philosophy and Probability colloquium series. Jeremy Steeger thanks audiences in the Controlled Quantum Dynamics group at Imperial College London. Our work on this project was supported by NSF Grant #1734155.

A Appendix

Before presenting the proofs of our results, we include the following comment on the relationship between Feintzeig and Fletcher’s *weak hidden variable representations* and our *WPS representations*. First, recall that Feintzeig and Fletcher directly construct representations of ‘quantum mechanical experiments’, viz. triples $(\mathcal{H}, \psi, \mathcal{O}_n)$ where \mathcal{H} is a Hilbert space, $\psi \in \mathcal{H}$ is a normalized vector, and $\mathcal{O}_n = \{P_1, \dots, P_n\}$ is a finite set of projectors on \mathcal{H} . We also note that in (Abramsky and Brandenburger 2011), an empirical model e is said to have a *quantum representation* if there is some $\mathcal{O}_n = X$ whose maximal subsets of commuting elements form the elements of \mathcal{M} , $O = \{0, 1\}$ (the spectrum of each P), and $e_C(s) = \langle \psi, P_s \psi \rangle$, where P_s is the projector on \mathcal{H} corresponding to the section $s \in \mathcal{E}(C)$. Thus, every quantum mechanical experiment is the quantum representation of the appropriate empirical model e .

In (Feintzeig and Fletcher 2017), a *weak hidden variable representation* of a quantum mechanical experiment $(\mathcal{H}, \psi, \mathcal{O}_n)$ is a weak probability space (Y, Σ, μ) and a map $E' : \mathcal{O}_n \rightarrow \Sigma$ satisfying the following two conditions: $\mu(E'(P)) = \langle \psi, P \psi \rangle$ for all $P \in \mathcal{O}_n$, and for all orthogonal $P_i, P_j \in \mathcal{O}_n$, $E'(P_i) \cap E'(P_j) \in \Sigma$ and $\mu(E'(P_i) \cap E'(P_j)) = 0$. In order to derive a connection between these representations and additivity-violations, Feintzeig and Fletcher require that these spaces satisfy an analog of (WC): for every set $Q \subseteq \mathcal{O}_n$ of mutually orthogonal projectors spanning \mathcal{H} , and letting Σ_Q be the algebra formed by Q , $\Sigma_Q \subseteq \Sigma$ and $(Y, \Sigma_Q, \mu|_Q)$ is a probability space.

Now let e be the empirical model induced by $(\mathcal{H}, \psi, \mathcal{O}_n)$. It is clear that by (WC), (EC), and (ME), any WPS representation e^\dagger will also be a weak hidden variable representation of $(\mathcal{H}, \psi, \mathcal{O}_n)$ where $E'(P) = E(P \mapsto 1)$ for all $P \in \mathcal{O}_n$. However, not every weak hidden variable representation of $(\mathcal{H}, \psi, \mathcal{O}_n)$ will be a WPS representation of e ; the former are blind to the structure of the event sheaf and so need not satisfy condition (3.5). We note that condition (3.5) imposes the minimal additional structure needed in order to derive interesting connections between sheaf-theoretic inconsistency and violations of additivity; these are established by Theorem 1.

We now present the proofs of this theorem and our two propositions. We begin by reviewing some notation. First, for AB-event sections, we will use a subscript to keep track of the set $U \subseteq X$ over which a section lies, i.e. $s_U \in \mathcal{E}(U)$ (where X is the measurement set of an empirical scenario). For instance, s_X will denote a particular global section. Second, we will use $S_U := \bar{E}(s_U)$ to simplify our notation for representations of AB-events (as sets within the WPS sample space Y); we will also use $\mathcal{E}^\dagger(U) := \{\bar{E}(s_U) : s_U \in \mathcal{E}(U)\}$ to denote the set of representations of AB-events over U . Third, in what follows, we will always use C and C' to refer to *maximal* contexts, i.e. $C, C' \in \mathcal{M}$.

In order to relate the different tiers of contextuality to properties of our WPS representations, we will want to excise the following two collections of sets from Y :

$$D_1 := \{E(s) \cap E(s') : s, s' \in \mathcal{E}(x) \text{ for some } x \in X\}. \quad (\text{A.1})$$

$$D_2 := \left\{ Y - \bigcup_{s \in \mathcal{E}(x)} E(s) : x \in X \right\}. \quad (\text{A.2})$$

D_1 represents ‘contradictory events’, whereas D_2 represents ‘measurements with non-existent outcomes’; neither of these types of ‘events’ is represented within the AB framework, and by (WC), (EC) and (ME), every set in these collections is a measure-zero Σ -event. Thus, it is the result of the excision $Z := Y - \bigcup (D_1 \cup D_2)$ which will contain the properties of interest to us. In the course of our analysis, we will frequently have reason to consider the intersection of Z with some set $B \in P(Y)$; we thus introduce the notation:

$$\tilde{B} := Z \cap B \quad (\text{A.3})$$

Let e be an empirical model, and let e^\dagger be any one of its WPS representations. We will now prove a key lemma that will be used repeatedly to help translate the ‘contextuality’ of e into the properties of e^\dagger .

Lemma 1. *If $z \in Z := Y - \bigcup (D_1 \cup D_2)$, then $z \in \tilde{E}(s_X)$ for some $s_X \in \mathcal{E}(X)$.*

Proof.

We begin by arguing that for all $x \in X$, any $z \in Z$ is contained in exactly one $S_x \in \mathcal{E}^\dagger(x)$. Let us define the union

$$A_x := \bigcup_{S_x \in \mathcal{E}^\dagger(x)} S_x$$

for an arbitrary $x \in X$. Since D_2 contains A_x^c and D_2 is disjoint from Z (by the definition of Z), it is clear that $Z \cap A_x^c = \emptyset$, and so $A_x \supseteq Z$. In other words, any $z \in Z$ is contained in the union A_x . To further see that any $z \in Z$ is contained in exactly one S_x out of this union, we simply note that since D_1 is disjoint from Z ,

$$\tilde{S}_x \cap \tilde{S}'_x = \emptyset \text{ for any } S, S' \in \mathcal{E}^\dagger(x), \quad (\text{A.4})$$

where we recall that $\tilde{S} := S \cap Z$.

We now know that for any $z \in Z$, there is a collection $\{S_x\}_{x \in X}$ such that each S_x contains z . Since any collection of AB-events $\{s_x\}_{x \in X}$ defines a global section s_X (by the sheaf property of \mathcal{E}), and the corresponding S_X is in the intersection of $\{S_x\}$ (by Eq. (3.5)), we conclude that $z \in Z$ is contained in S_X . \square

We now use this lemma to prove Theorem 1.

Proof of Theorem 1.

(1) *e strongly contextual $\implies e^\dagger$ maximally violates subadditivity*

In order to show the existence of a collection of sets that maximally violate subadditivity, it will be convenient to define D_3 , the collection of measure-zero sets S_C each of which contains some representation S_X of a global section. More formally:

$$D_3 := \{S_C \text{ corresponding to some } s_C : \mu(S_C) = 0 \text{ and } S_C \supset S_X \text{ corresponding to some } s_X\}.$$

Our strategy will be to use strong contextuality to show that the union of sets in D_3 contains Z , thus providing—in conjunction with D_1 and D_2 —the collection of sets which maximally violates subadditivity.

We now translate ‘strong contextuality’ into a feature of the collection D_3 . Recall that strong contextuality means that there are no global sections in the support of e ; in other words, for any global section s_X , there exists a section $s_C := s_X|_C$ such that $e_C(s_C) = 0$. Thus, by (EC) and condition (3.5), strong contextuality implies that any S_X is contained in some measure-zero S_C ; hence such an S_C belongs in D_3 . We can now combine this fact with Lemma 1 to deduce that—since any $z \in Z$ is contained in some S_X —strong contextuality implies that $\bigcup D_3 \supset Z$.

In light of this last WPS characterization of strong contextuality, we will take our violation-inducing collection to be $\mathcal{V}_3 := D_1 \cup D_2 \cup D_3$. Since $Y = \bigcup \mathcal{V}_3$ and $\mu(Y) = 1$ by (WC), we can immediately compute that $\mathfrak{a}(\mathcal{V}_3) = \mu(\bigcup \mathcal{V}_3) - \sum_{A \in \mathcal{V}_3} \mu(A) = 1$, i.e. \mathcal{V}_3 yields a maximal violation of subadditivity.

(2) e logically contextual $\implies e^\dagger$ violates subadditivity

Our strategy here is similar, viz. we will find a collection of sets that covers Y and use this to show that the collection violates subadditivity. However, our previous argument will need to be supplemented, because logical contextuality only guarantees the existence of *one* section $s^* \in \mathcal{E}(C)$ in the support of e which does not extend to a global section in the support of e . Thus, we will only be able to show that the union of sets in D_3 contains a subset of Z ; in order to construct a collection that completely contains Z , we will need to introduce a new collection that we call D_4 .

From the definition of logical contextuality, we know that for all global sections s_X that restrict to s^* , there exists a distinct maximal context C' such that the restriction $s_X|_{C'} = s_{C'}$ is not in the support of $e_{C'}$, i.e. $e_{C'}(s_{C'}) = 0$. Thus, by (Empirical Consistency) and condition (3.5), $S_{C'}$ is a measure-zero set that contains any such S_X , i.e. it is an element of D_3 .

Combining this fact with (3.5) and Lemma 1 now lets us argue that $\bigcup D_3$ contains $\tilde{S}^* := S^* \cap Z$. To see this, we first note that by (3.5) and Lemma 1, $S^* = \bigcup_{S_X: S_X|_C = S^*} S_X$. And since we have just argued that any S_X that extends to S^* is contained in $S_{C'}$, it follows that $\bigcup D_3 \supset \tilde{S}^*$.

However, to complete our argument, we will need to generate a collection whose union contains all of Z . In order to do so, we first define another collection $D_4 := \{S_C : S_C \neq S^*\}$ (i.e. sets that represent sections over C which are not s^*), and note that by (EC), any set in D_4 must have measure less than one because—by hypothesis— $\mu(S^*) \neq 0$. Supplementing D_3 with D_4 then allows us to construct the desired collection: $\bigcup(D_3 \cup D_4) \supset Z$, because by Lemma 1 and (3.5), any $z \in Z$ is contained in precisely one S_C .

We can now produce our subadditivity collection by defining $\mathcal{V}_4 := D_1 \cup D_2 \cup D_3 \cup D_4$, and noting that $Y = \bigcup \mathcal{V}_4$. Since the sets in D_1, D_2, D_3 are all of measure zero and the sets in D_4 are of measure less than one, it follows that $\alpha(\mathcal{V}_4) = \mu(\bigcup \mathcal{V}_4) - \sum_{A \in \mathcal{V}_4} \mu(A) > 0$. Thus, \mathcal{V}_4 violates subadditivity.

(3) e probabilistically contextual \implies any monotonic extension of e^\dagger violates additivity

Note that if e is possibilistically (i.e. either strongly or logically) contextual, then any monotonic extension of e^\dagger will violate subadditivity, and so it will also violate additivity. Thus, we will only need to consider the case in which e is probabilistically, but not possibilistically, contextual. In this case, e^\dagger satisfies subadditivity, and its monotonic extensions which are *not* subadditive are not additive by default; thus, we need only consider the monotonic extensions of e^\dagger which satisfy subadditivity. As such, we will now argue that if e is probabilistically contextual, then any subadditive and monotonic extension of e^\dagger will contain a collection of sets \mathcal{V} which violates additivity.

By probabilistic contextuality and (EC), there must exist some $S^* \in \mathcal{E}^\dagger(C)$ such that $\mu(S^*)$ cannot be recovered by marginalization from any μ' , i.e.

$$\mu(S^*) \neq \sum_{\{S_X \in \mathcal{E}^\dagger(C) : S_X|_C = S^*\}} \mu'(S_X). \quad (\text{A.5})$$

The non-marginalization equation (A.5) bears a resemblance to the additivity-violation condition; however, there is insufficient information to show that it is defined on a disjoint collection $\{S_X\}$ whose union yields S^* . In order to obtain a related collection of sets that satisfies these two properties, we define $\mathcal{V} := \{\tilde{S}_X : S_X|_C = S^*\}$. We then observe that by Lemma 1, this is a disjoint collection; and that by Lemma 1 and (3.5), $\bigcup \mathcal{V} = \tilde{S}^*$, which relates the collection to the LHS of (A.5).

In order to relate the measures of the sets used in (A.5) to the measures of the sets in \mathcal{V} , we will need to establish two further facts, viz. $\mu'(\tilde{S}_X) = \mu'(S_X)$ and $\mu'(\bigcup \mathcal{V}) = \mu'(\tilde{S}^*) = \mu(S^*)$. We do so by arguing that $\mu'(\tilde{A}) = \mu'(A)$ for any $A \in \Sigma'$. Let $\tilde{A} := A \cap Z^c$. To obtain the result, note that $\mu'(Z^c) = 0$ by the subadditivity of μ' and so $\mu'(\tilde{A}) = 0$ by the monotonicity of μ' . Then observe that by subadditivity, $\mu'(A) = \mu'(\tilde{A} \cup \tilde{A}) \leq \mu'(\tilde{A}) + \mu'(\tilde{A}) \leq \mu'(\tilde{A})$, and again by monotonicity, $\mu'(\tilde{A}) \leq \mu'(A)$. Thus the result follows.

By means of the above facts, we can now use (A.5) to compute that $\mu'(\bigcup \mathcal{V}) \neq \sum_{A \in \mathcal{V}} \mu'(A)$, thus showing that \mathcal{V} witnesses an additivity-violation for (Y, Σ', μ') . □

In preparation for the proofs of Propositions 1 and 2, we provide an illustration of a simple combinatorial WPS representation. Let e be the empirical model of the Specker triangle, where $X = \{a, b, c\}$, $O = \{0, 1\}$, and $\mathcal{M} = \{\{a, b\}, \{b, c\}, \{a, c\}\}$.²² A simple combinatorial WPS representation e^\dagger is obtained by letting the sample space Y contain a unique point for each global section $s_X \in \mathcal{E}(X)$ and no additional points.²³ This WPS can be graphically represented by the cube in Figure A.1. In this figure, the points exhaust Y , edges represent events over maximal contexts, and faces represent atomic events. For example, the shaded face is $\bar{E}(c \mapsto 0) = \{a_0b_0c_0, a_0b_1c_0, a_1b_0c_0, a_1b_1c_0\}$, and the thick edges represent the set $\{S : S \in \mathcal{E}^\dagger(\{a, b\})\}$.²⁴

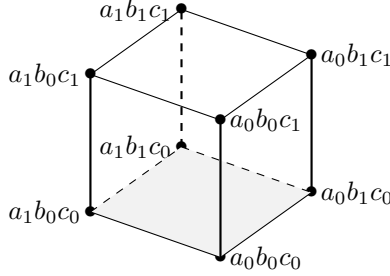


Figure A.1: A combinatorial WPS-representation of the Specker triangle.

Before proving Proposition 1, we recall that $\mathcal{V}_{\mathcal{M}} := \{S_C : S_C \in \mathcal{E}^\dagger(C), C \in \mathcal{M}\}$ and that an ‘additive cover’ is collection of disjoint sets $A \subset \Sigma$ that covers the sample space such that $\mathfrak{a}(A) = 0$.

Proof of Proposition 1.

(1) e strongly contextual $\iff e^\dagger$ strongly $\mathcal{V}_{\mathcal{M}}$ -violates subadditivity

Recall that e^\dagger strongly $\mathcal{V}_{\mathcal{M}}$ -violates subadditivity iff the collection V of measure-zero sets in $\mathcal{V}_{\mathcal{M}}$ covers all the measure-nonzero elements of an additive cover $A \subseteq \mathcal{V}$. By (Exhaustiveness), we can set $A = \{S_C : S_C \in \mathcal{E}^\dagger(C)\}$ for some $C \in \mathcal{M}$. It is clear that e^\dagger strongly $\mathcal{V}_{\mathcal{M}}$ -violates subadditivity iff $\bigcup V = Y$. We will now show that this holds iff e is strongly contextual.

Recall that strong contextuality means that there are no global sections in the support of e ; in other words, for any global section s_X , there exists a section $s_C := s_X|_C$ such that $e_C(s_C) = 0$.

Thus, by (EC), (3.5), and (Strong Mutual Exclusivity), every S_X is contained in some measure-zero S_C iff e is strongly contextual. By (Exhaustiveness), $\bigcup_{S_X \in \mathcal{E}^\dagger(X)} S_X = Y$, and by condition (3.5), no S_X is trivial. Thus, $\bigcup V = Y$ iff e is strongly contextual.

(2) e logically contextual $\iff e^\dagger$ logically $\mathcal{V}_{\mathcal{M}}$ -violates subadditivity

Recall that e^\dagger logically \mathcal{V} -violates subadditivity iff the collection V of measure-zero sets in $\mathcal{V} \subseteq \Sigma$ covers some non-measure-zero element a of an additive cover $A \subseteq \mathcal{V}$ (and so $\mathfrak{a}(V \cup A \setminus \{a\}) > 0$).

Suppose e is logically contextual; this means there is some section $s^* \in \mathcal{E}(C)$ for some $C \in \mathcal{M}$ which is in the support of e but which does not extend to a global section in the support of e . By condition (3.5), for each global section s_X that restricts to s^* , the associated WPS-event S_X is contained in some $S_C \in V$. Again by (3.5), the union of all these S_X is equal to $E(s^*) = S^*$; thus $\bigcup V$ contains S^* . Let $A = \{S_C : S_C \in \mathcal{E}^\dagger(C)\}$ and note that A covers Y (by Exhaustiveness) and contains S^* . Thus, $(A \setminus \{S^*\}) \cup V$ yields the desired logical violation of subadditivity, viz. $\mathfrak{a}(V \cup A \setminus \{S^*\}) > 0$.

²²This is all the data we will need for our example, but see (Liang et al. 2011) for the complete model.

²³Note that the recipe for this simple WPS representation recovers the formalism of ‘experimental behaviors’ used by (Dowker et al. 2014).

²⁴Feintzeig and Fletcher (2017) discuss a variety of types of weak probability spaces that satisfy certain additional conditions, but not all of these will admit of such combinatorial models. The model illustrated, for example, fails to be a Generalized Probability Space (GPS) representation of the Specker triangle because there are disjoint Σ -events whose union is not a Σ -event.

Now suppose e is not logically contextual; this means *all* sections s_C in the support of e extend to global sections s_X in the support of e . By (3.5), any cover of a measure-nonzero S_C must cover all of the global sections to which it extends; that is, it must cover all S_X such that $S_X|_C = S_C$. Since at least one of these S_X is in the support of e^\dagger , by (EC) and (Strong Mutual Exclusivity), any cover using elements of \mathcal{V}_M must contain some measure-nonzero element. Thus no logical violation of subadditivity can be constructed.

(3) e probabilistically contextual $\iff e^\dagger \mathcal{V}_M$ -violates additivity

Recall $e^\dagger \mathcal{V}_M$ -violates additivity iff the algebra generated by \mathcal{V}_M contains some collection V of disjoint sets that violates additivity in any monotonic extension of e^\dagger (and so $\mathfrak{a}(V) \neq 0$ in these extensions).

First, note that $\mathcal{E}^\dagger(X)$ is a collection of disjoint sets that generates the algebra Σ' that is generated by \mathcal{V}_M . Now note that if e is *not* probabilistically contextual, then (by definition) there exists some $e_X \in \mathcal{D}\mathcal{E}(X)$. This e_X induces a monotonic extension (Y, Σ', μ') such that, by (EC) and (WC), all elements in Σ' marginalize appropriately. That is: for any $S \in U$ for any $U \subseteq X$,

$$\mu(S) = \sum_{\{S_X \in \mathcal{E}^\dagger(X) : S_X|_U = S\}} \mu'(S_X). \quad (\text{A.6})$$

Since the sets in $\mathcal{E}^\dagger(X)$ are all disjoint and generate the algebra, there is no collection of disjoint sets in this algebra that yield a violation of additivity. So there is an additive, monotonic extension of e^\dagger .

Now suppose that e is probabilistically contextual. By (EC), in any monotonic extension (Y, Σ', μ') , there must exist some $S^* \in \mathcal{E}^\dagger(C)$ such that $\mu(S^*)$ cannot be recovered by marginalization from μ' . Let $V = \{S_X \in \mathcal{E}^\dagger(X) : S_X|_C = S^*\}$, and note that V is a collection of disjoint sets such that $\mathfrak{a}(V) \neq 0$.²⁵ \square

We now turn to Proposition 2. Its proof makes use of the following consequence of condition (3.5), (Strong Mutual Exclusivity), and (Exhaustiveness): letting $\mathbb{V}|_{\mathcal{V}_M} := \{V|_{\mathcal{V}_M} : V \in \mathbb{V}\}$, there is a bijection

$$v : \mathcal{E}(X) \rightarrow \mathbb{V}|_{\mathcal{V}_M} :: s_X \mapsto V^{s_X}, \quad (\text{A.7})$$

where for $C \in \mathcal{M}$, $V^{s_X}(S_C)$ is 1 if $S_X|_C = S_C$ and 0 otherwise. Additionally, if we let $\mathcal{D}_A\mathcal{E}(X)$ be the presheaf containing *all* the distributions on X , then there is a bijection

$$d : \mathcal{D}_A\mathcal{E}(X) \rightarrow \text{Conv}(\mathbb{V}|_{\mathcal{V}_M}) :: e_X \mapsto \sum_{s_X \in \mathcal{E}(X)} e_X(s_X) \cdot v(s_X). \quad (\text{A.8})$$

By means of these two maps, we can demonstrate the equivalence of the hierarchy of contextuality and the hierarchy of convexity-violation.

Proof of Proposition 2.

(1) e strongly contextual $\iff e^\dagger$ strongly \mathcal{V}_M -violates \mathbb{V} -convexity

(\implies) Suppose towards a contradiction that (4.3) is satisfied for some \mathbb{V}' and some f . Pick a $V \in \mathbb{V}'$; then $v^{-1}(V|_{\mathcal{V}_M}) = s_X$ must be in the support of e , which yields a contradiction.

(\impliedby) Suppose towards a contradiction that $s_X \in \mathcal{E}(X)$ is in the support of e . But then the set of $V \in \mathbb{V}$ which restricts to $v(s_X)$ satisfies (4.3) for an appropriate choice of f , which yields a contradiction.

(2) e logically contextual $\iff e^\dagger$ logically \mathcal{V}_M -violates \mathbb{V} -convexity

(\implies) Suppose that $\chi_\mu|_{\mathcal{V}_M}$ is the logical sum of the elements of $\mathbb{V}' \subseteq \mathbb{V}$. Then every local section in the support of e extends to a global section in the pre-image of v over \mathbb{V}' , which yields a contradiction.

(\impliedby) Suppose that all local sections in the support of e extend to global sections; then the logical sum of the image of v over these global sections yields $\chi_\mu|_{\mathcal{V}_M}$, which yields a contradiction.

²⁵These considerations are sufficient to demonstrate the stronger result that e is probabilistically contextual if and only if *any* (monotonic or non-monotonic) extension of e^\dagger contains a subset in Σ' that violates additivity. We restrict our focus to monotonic extensions to emphasize the connection to Theorem 1.

(3) e probabilistically contextual $\iff e^\dagger \mathcal{V}_\mathcal{M}$ -violates \mathbb{V} -convexity

(\implies) Suppose otherwise, viz. suppose that some convex combination of \mathbb{V} yields $\mu|_{\mathcal{V}_\mathcal{M}}$. It follows that $d^{-1}(\mu|_{\mathcal{V}_\mathcal{M}})$ is a distribution e_X that marginalizes to each e_C , which yields a contradiction.

(\impliedby) Suppose otherwise, viz. suppose that there is some $e_X \in \mathcal{DE}(X)$. Then $d(e_X)$ yields an element of $\text{Conv}(\mathbb{V})$ which is equal to $\mu|_{\mathcal{V}_\mathcal{M}}$; contradiction. □

References

- Abramsky, Samson and Adam Brandenburger (2011), “The sheaf-theoretic structure of non-locality and contextuality”, *New Journal of Physics*, 13, 11, p. 113036.
- Acín, A., T. Fritz, A. Leverrier, and A. Belén Sainz (2012), “A Combinatorial Approach to Nonlocality and Contextuality”, *ArXiv e-prints* (Dec. 2012), arXiv: 1212.4084 [quant-ph].
- Bell, John S. (1964), “On the Einstein-Podolsky-Rosen Paradox”, *Physics*, 1, pp. 195-200.
- (1966), “On the Problem of Hidden Variables in Quantum Mechanics”, *Reviews of Modern Physics*, 38, 3, pp. 447-452.
- Bradley, Seamus (2016), “Nonclassical Probability and Convex Hulls”, *Erkenntnis*, 41, pp. 811-840.
- Cabello, Adán, José M. Estebarez, and Guillermo García-Alcaine (1996), “Bell-Kochen-Specker theorem: a proof with 18 vectors”, *Physics Letters A*, 212, 4, pp. 183-187.
- Craig, David, Fay Dowker, Joe Henson, Seth Major, David Rideout, and Rafael D Sorkin (2006), “A Bell inequality analog in quantum measure theory”, *Journal of Physics A: Mathematical and Theoretical*, 40, 3, p. 501.
- De Finetti, Bruno (1937), “La prévision: ses lois logiques, ses sources subjectives”, *Annales de l’Institut Henri Poincaré*, 17, pp. 1-68; trans. “Foresight: Its Logical Laws, Its Subjective Sources”, in *Studies in Subjective Probability*, trans. by Henry E. Kyburg, Jr., 2nd ed., Robert E. Krieger, Huntington 1980.
- De Finetti, Bruno (1974), *Theory of Probability*, trans. by Antonio Machi and Adrian Smith, 1st ed., Wiley, New York, vol. 1.
- Dowker, Fay, Joe Henson, and Petros Wallden (2014), “A histories perspective on characterizing quantum non-locality”, *New Journal of Physics*, 16, 3, p. 033033.
- Feintzeig, Benjamin (2015), “Hidden Variables and Incompatible Observables in Quantum Mechanics”, *The British Journal for the Philosophy of Science*, 66, pp. 905-927.
- Feintzeig, Benjamin and Samuel C. Fletcher (2017), “On Noncontextual, Non-Kolmogorovian Hidden Variable Theories”, *Foundations of Physics*, 47, pp. 294-315.
- Fine, Arthur (1982a), “Hidden Variables, Joint Probability, and the Bell Inequalities”, *Physical Review Letters*, 48, 5, pp. 291-295.
- (1982b), “Joint distributions, quantum correlations, and commuting observables”, *Journal of Mathematical Physics*, 23, pp. 1306-1310.
- Gudder, Stanley P (1973), “Generalized Measure Theory”, *Foundations of Physics*, 3, 3, pp. 399-411.
- Hájek, Alan (2008), “Dutch Book Arguments”, in *The Oxford Handbook of Rational and Social Choice*, ed. by Paul Anand, Prasanta Pattanaik, and Clemens Puppe, Oxford University Press.
- Hardy, Lucien (1993), “Nonlocality for two particles without inequalities for almost all entangled states”, *Physical Review Letters*, 71, 11, p. 1665.
- Kaplan, Mark (1998), *Decision theory as philosophy*, Cambridge University Press.
- Kochen, Simon and E. P. Specker (1975), “The Problem of Hidden Variables in Quantum Mechanics”, in *The Logico-Algebraic Approach to Quantum Mechanics: Volume I: Historical Evolution*, Springer, pp. 293-328.
- Kühr, Jan and Daniele Mundici (2007), “De Finetti theorem and Borel states in $[0,1]$ -valued algebraic logic”, *International Journal of Approximate Reasoning*, 46, 3, pp. 605-616.
- Kunjwal, Ravi (2015), “Fine’s theorem, noncontextuality, and correlations in Specker’s scenario”, *Physical Review A*, 91, 2, p. 022108.
- Liang, Yeong-Cherng, Robert W. Spekkens, and Howard M. Wiseman (2011), “Specker’s parable of the overprotective seer: A road to contextuality, nonlocality and complementarity”, *Physics Reports*, 506, 1, pp. 1-39.

- Murofushi, Toshiaki and Michio Sugeno (1989), “An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure”, *Fuzzy Sets and Systems*, 29, 2, pp. 201-227.
- Paris, J. B. (2001), “A note on the Dutch Book method”, in *Second International Symposium on Imprecise Probabilities and their Applications*, <http://www.sipta.org/isipta01/proceedings/>.
- Penrose, Roger (1992), “On the cohomology of impossible figures”, *Leonardo*, 25, 3, pp. 245-247.
- Pitowsky, Itamar (1989), *Quantum Probability—Quantum Logic*, Springer-Verlag, Berlin.
- Popescu, Sandu (2014), “Nonlocality beyond quantum mechanics”, *Nature Physics*, 10, 4, pp. 264-270.
- Roumen, F. (2016), “Cohomology of Effect Algebras”, *ArXiv e-prints* (Feb. 2016), arXiv: 1602.00567 [math.QA].
- Spekkens, Robert W. (2005), “Contextuality for Preparations, Transformations, and Unsharp Measurements”, *Physical Review A*, 71, 5, p. 052108.
- Staton, Sam and Sander Uijlen (2015), “Effect Algebras, Presheaves, Non-locality and Contextuality”, in *Proceedings, Part II, of the 42nd International Colloquium on Automata, Languages, and Programming*, ed. by M. Halldórsson, K. Iwama, N. Kobayashi, and B. Speckmann, Springer-Verlag, vol. 9135, pp. 401-413.
- Steeger, Jeremy (2018), “Probabilism for Stochastic Theories”, *PhilSci Archive e-prints*, <http://philsci-archive.pitt.edu/15014/>.
- Suppes, Patrick and Mario Zanotti (1991), “Existence of hidden variables having only upper probabilities”, *Foundations of Physics*, 21, 12, pp. 1479-1499.
- van den Berg, B. and C. Heunen (2010), “Noncommutativity as a colimit”, *ArXiv e-prints* (Mar. 2010), arXiv: 1003.3618 [math.CT].
- Williams, J. Robert G. (2012), “Generalized Probabilism: Dutch Books and Accuracy Domination”, *Journal of Philosophical Logic*, 41, pp. 811-840.
- Wroński, Leszek and Michał Tomasz Godziszewski (2017), “Dutch Books and nonclassical probability spaces”, *European Journal for Philosophy of Science*, 7, pp. 267-284.